Approximating Fixation Probabilities in the Generalized Moran Process

Josep Díaz,* Leslie Ann Goldberg,† George B. Mertzios,‡
David Richerby,† Maria Serna* and Paul G. Spirakis§

Abstract

We consider the Moran process, as generalized by Lieberman, Hauert and Nowak (*Nature*, 433:312–316, 2005). A population resides on the vertices of a finite, connected, undirected graph and, at each time step, an individual is chosen at random with probability proportional to its assigned "fitness" value. It reproduces, placing a copy of itself on a neighbouring vertex chosen uniformly at random, replacing the individual that was there. The initial population consists of a single mutant of fitness $r > 0$ placed uniformly at random, with every other vertex occupied by an individual of fitness 1. The main quantities of interest are the probabilities that the descendants of the initial mutant come to occupy the whole graph (fixation) and that they die out (extinction); almost surely, these are the only possibilities. In general, exact computation of these quantities by standard Markov chain techniques requires solving a system of linear equations of size exponential in the order of the graph so is not feasible. We show that, with high probability, the number of steps needed to reach fixation or extinction is bounded by a polynomial in the number of vertices in the graph. This bound allows us to construct fully polynomial randomized approximation schemes (FPRAS) for the probability of fixation (when $r \geqslant 1$) and of extinction (for all $r > 0$).

Keywords: Evolutionary dynamics, Markov-chain Monte Carlo, approximation algorithm.

1 Introduction

Population and evolutionary dynamics have been extensively studied [2, 7, 16, 27, 30–32], usually with the assumption that the evolving population has no spatial structure. One of the main models in this area is the Moran process [23]. The initial population contains a single "mutant" with fitness $r > 0$, with all other individuals having fitness 1. At each step of the process, an individual is chosen at random, with probability proportional to its fitness. This individual reproduces, replacing a second individual, chosen uniformly at random, with a copy of itself. Population dynamics has also been studied in the context of strategic interaction in evolutionary game theory [10, 13–15, 29].

Lieberman, Hauert and Nowak introduced a generalization of the Moran process, where the members of the population are placed on the vertices of a connected graph which is, in general, directed [19, 26]. In this model, the initial population again consists of a single mutant of fitness $r > 0$ placed on a vertex chosen uniformly at random, with each other vertex occupied by a non-mutant with fitness 1. The individual that will reproduce is chosen as before but now one of its neighbours is randomly selected for replacement, either uniformly or according to a weighting of the edges. The original Moran process can be recovered by taking the graph to be an unweighted clique. In the present paper, we consider the process on finite, unweighted, undirected graphs.

Several similar models describing particle interactions have been studied previously, including the SIR and SIS epidemic models [9, Chapter 21], the voter model, the antivoter model and the exclusion process [1, 8, 20]. Related models, such as the decreasing cascade model [18, 24], have been studied in the context of influence propagation in social networks and other models have been considered for dynamic monopolies [3]. However, these models do not consider different fitnesses for the individuals.

In general, the Moran process on a connected, directed graph may end with all vertices occupied by mutants or with no vertex occupied by a mutant — these cases are referred to as *fixation* and *extinction*, respectively — or the process may continue forever. However, for undirected graphs and strongly-connected digraphs, the process terminates almost surely, either at

fixation or extinction. (We consider only finite graphs.) At the other extreme, fixation is impossible in the directed graph with vertices $\{x, y, z\}$ and edges $\{\overrightarrow{xz}, \overrightarrow{yz}\}$ and extinction is impossible unless the mutant starts at z. The *fixation probability* for a mutant of fitness r in a graph G is the probability that fixation is reached and is denoted $f_{G,r}$.

The fixation probability can be determined by standard Markov chain techniques. However, doing so for a general graph on n vertices requires solving a set of 2^n linear equations, which is not computationally feasible. As a result, most prior work on computing fixation probabilities in the generalized Moran process has either been restricted to small graphs [7] or graph classes where a high degree of symmetry reduces the size of the set of equations — for example, paths, cycles, stars and cliques [4–6] — or has concentrated on finding graph classes that either encourage or suppress the spread of the mutants [19, 22]. Rychtář and Stadler present some experimental results on fixation probabilities for random graphs derived from grids [28].

Because of the apparent intractability of exact computation, we turn to approximation. Using a potential function argument, we show that, with high probability, the Moran process on an undirected graph of order n reaches absorption (either fixation or extinction) within $\mathcal{O}(n^5)$ steps if $r = 1$ and $\mathcal{O}(n^4)$ and $\mathcal{O}(n^3)$ steps when $r > 1$ and $r < 1$, respectively. Taylor et al. [31] studied absorption times for variants of the generalized Moran process but, in our setting, their results only apply to the process on regular graphs, where it is equivalent to a biased random walk on a line with absorbing barriers. The absorption time analysis of Broom et al. [4] is also restricted to cliques, cycles and stars. In contrast to this earlier work, our results apply to all connected undirected graphs.

Our bound on the absorption time, along with polynomial upper and lower bounds for the fixation probability, allows the estimation of the fixation and extinction probabilities by Monte Carlo techniques. Specifically, we give a *fully polynomial randomized approximation scheme* (FPRAS) for these quantities. An FPRAS for a function $f(X)$ is a polynomial-time randomized algorithm g that, given input X and an error bound ε satisfies $(1 - \varepsilon)f(X) \leqslant g(X) \leqslant (1 + \varepsilon)f(X)$ with probability at least $\frac{3}{4}$ and runs in time polynomial in the length of X and $\frac{1}{\varepsilon}$ [17].

For the case $r < 1$, there is no positive polynomial lower bound on the fixation probability so only the extinction probability can be approximated by this technique. (Note that, when $f \ll 1$, computing $1 - f$ to within a factor of $1 \pm \varepsilon$ does not imply computing f to within the same factor.)

Notation. Throughout, we consider only finite, connected, undirected graphs $G = (V, E)$ and we write $n = |V|$ (the *order* of the graph). Our results apply only to connected graphs as, for any disconnected graph, the fixation probability is necessarily zero; we also exclude the one-vertex graph to avoid trivialities. The edge between vertices x and y is denoted by xy. For a subset $X \subseteq V(G)$, we write $X + y$ and $X - y$ for $X \cup \{y\}$ and $X \setminus \{y\}$, respectively.

Throughout, r denotes the fitness of the initially introduced mutant in the graph. Given a set $X \subseteq V$, we denote by $W(X) = r|X| + |V \setminus X|$ the *total fitness* of the population when exactly the vertices of X are occupied by mutants. We write $f_{G,r}(x)$ for the fixation probability of G, given that the initial mutant with fitness r was introduced at vertex x. We denote by $f_{G,r} = \frac{1}{n} \sum_{x \in V} f_{G,r}(x)$ the *fixation probability of* G; that is, the probability that a single mutant with fitness r placed uniformly at random in V eventually takes over the graph G. Finally, we define the problem MORAN FIXATION (respectively, MORAN EXTINCTION) as follows: given a graph $G = (V, E)$ and a fitness value $r > 0$, compute the value $f_{G,r}$ (respectively, $1 - f_{G,r}$).

Organization of the paper. In Section 2, we demonstrate polynomial upper and lower bounds for the fixation probability $f_{G,r}$ for an arbitrary undirected graph G. In Section 3, we use our potential function to derive polynomial bounds on the absorption time (both in expectation and with high probability) in general undirected graphs. Our FPRAS for computing fixation and extinction probabilities appears in Section 4.

2 Bounding the fixation probability

Lieberman et al. observed that, if G is a directed graph with a single source (a vertex with in-degree zero), then $f_{G,r} = \frac{1}{n}$ for any value of fitness $r > 0$ of the initially introduced mutant [19] (see also [26, p. 135]). In the following lemma we prove that $f_{G,r} \geqslant \frac{1}{n}$ for every undirected graph, whenever $r \geqslant 1$.

LEMMA 2.1. *Let* $G = (V, E)$ *be an undirected graph with* n *vertices. Then* $f_{G,r} \geqslant \frac{1}{n}$ *for any* $r \geqslant 1$.

Proof. Consider the variant of the process where every vertex starts with its own colour and every vertex has fitness 1. Allow the process to evolve as usual: at each step, a vertex is chosen uniformly at random and its colour is propagated to a neighbour also chosen uniformly at random. At any time, we can consider the vertices of any one colour to be the mutants and all the other vertices to be non-mutants. Hence, with probability 1, some colour will take over the graph and the probability that x's initial colour takes over is exactly $f_{G,1}(x)$. Thus, $f_{G,r} \geqslant f_{G,1} = \frac{1}{n} \sum_{x \in V} f_{G,1}(x) = \frac{1}{n}$. $\square$

Note that there is no corresponding polynomial lower bound when $r < 1$. For example, for $r \neq 1$, the fixation probability of the clique K_n is given by

$$f_{K_n,r} = \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^n}} \, .$$

For $r > 1$, this is at least $1 - \frac{1}{r}$ but there is no non-trivial polynomial lower bound where $r < 1$.

LEMMA 2.2. *Let $G = (V, E)$ be an undirected graph with n vertices. Then $f_{G,r} \leqslant 1 - \frac{1}{n+r}$ for any $r > 0$.*

Proof. For any vertex $x \in V$, let $Q(x) = \sum_{xy \in E} \frac{1}{\deg y}$. Note that $\sum_{x \in V} Q(x) = n$.

To give an upper bound for $f_{G,r}(x)$ for every $x \in V$, we relax the Markov chain by assuming that fixation is reached as soon as a second mutant is created. From the state $S = \{x\}$, the probability that a new mutant is created is $a(x) = \frac{r}{n-1+r}$ and the probability that one of x's non-mutant neighbours reproduces into x is $b(x) = \frac{1}{n-1+r}Q(x)$. The probability that the population stays the same, because a non-mutant reproduces to a non-mutant vertex, is $1 - a(x) - b(x)$. The probability that the mutant population reaches two (i.e., that the first change to the state is the creation of a new mutant) is given by

$$p(x) = \frac{a(x)}{a(x) + b(x)} = \frac{r}{r + Q(x)} \, .$$

Therefore, the probability that the new process reaches fixation is

$$p = \frac{1}{n} \sum_{x \in V} p(x) = \frac{r}{n} \sum_{x \in V} \frac{1}{r + Q(x)} \, .$$

Writing $p = \frac{r}{n} \sum_{i=1}^{n} (r + q_i)^{-1}$, we wish to find the maximum value subject to the constraints that $q_i > 0$ for all i and $\sum_{i=1}^{n} q_i = \sum_{x \in V} Q(x) = n$. If we relax the first constraint to $q_i \geqslant 0$, the sum is maximized by setting $q_1 = n$ and $q_2 = \cdots = q_n = 0$. Therefore,

$$f_{G,r} \leqslant p \leqslant \frac{r}{n}\left(\frac{1}{r+n} + (n-1)\frac{1}{r+0} \right) = 1 - \frac{1}{r+n} \, . \ \square$$

3 Bounding the absorption time

In this section, we show that the Moran process on a connected graph G of order n is expected to reach absorption in a polynomial number of steps. To do this, we use the potential function given by

$$\phi(X) = \sum_{x \in X} \frac{1}{\deg x}$$

for any state $X \subseteq V(G)$ and we write $\phi(G)$ for $\phi(V(G))$. Note that $1 < \phi(G) \leqslant n$ and that, if $(X_i)_{i \geqslant 0}$ is a Moran process on G then $\phi(X_0) = 1/\deg x \leqslant 1$ for some vertex $x \in V$ (the initial mutant).

First, we show that the potential strictly increases in expectation when $r > 1$ and strictly decreases in expectation when $r < 1$.

LEMMA 3.1. *Let $(X_i)_{i \geqslant 0}$ be a Moran process on a graph $G = (V, E)$ and let $\emptyset \subset S \subset V$. If $r \geqslant 1$, then*

$$\mathbb{E}[\phi(X_{i+1}) - \phi(X_i) \mid X_i = S] > \left(1 - \frac{1}{r}\right) \cdot \frac{1}{n^3} \, .$$

Otherwise,

$$\mathbb{E}[\phi(X_{i+1}) - \phi(X_i) \mid X_i = S] < \frac{r-1}{n^3} \, .$$

Proof. Write $W(S) = n + (r-1)|S|$ for the total fitness of the population. For $\emptyset \subset S \subset V$, and any value of r, we have

$$\mathbb{E}[\phi(X_{i+1}) - \phi(X_i) \mid X_i = S]$$
$$= \frac{1}{W(S)} \sum_{\substack{xy \in E \\ x \in S, y \in \overline{S}}} \left(r \cdot \frac{\phi(S+y) - \phi(S)}{\deg x} \right.$$
$$\left. + \frac{\phi(S-x) - \phi(S)}{\deg y} \right)$$
$$= \frac{1}{W(S)} \sum_{\substack{xy \in E \\ x \in S, y \in \overline{S}}} \left(r \cdot \frac{1}{\deg y} \cdot \frac{1}{\deg x} \right.$$
$$\left. - \frac{1}{\deg x} \cdot \frac{1}{\deg y} \right)$$
$$(3.1) \qquad = \frac{r-1}{W(S)} \sum_{\substack{xy \in E \\ x \in S, y \in \overline{S}}} \frac{1}{\deg x \deg y} \, .$$

The sum is minimized by noting that there must be at least one edge between S and $\overline{S}$ and that its endpoints have degree at most $(n-1) < n$. The greatest weight configuration is the one with all mutants if $r \geqslant 1$ and the one with no mutants if $r < 1$. Therefore, if $r \geqslant 1$, we have

$$\mathbb{E}[\phi(X_{i+1}) - \phi(X_i) \mid X_i = S] > \frac{r-1}{rn} \cdot \frac{1}{n^2}$$
$$= \left(1 - \frac{1}{r}\right)\frac{1}{n^3}$$

and, if $r < 1$,

$$\mathbb{E}[\phi(X_{i+1}) - \phi(X_i) \mid X_i = S] < (r-1)\frac{1}{n^3} \, . \ \square$$

The method of bounding in the above proof appears somewhat crude — for example, in a graph of order $n > 2$, if both endpoints of the chosen edge from S to $\overline{S}$ have degree $n-1$ then there must be more edges between mutants and non-mutants. Nonetheless, over the class of all graphs, the bound of Lemma 3.1 is asymptotically optimal up to constant factors. For $n \geqslant 2$, let G_n be the n-vertex graph made by adding an edge between the centres of two disjoint stars of as close-to-equal size as possible. If S is the vertex set of one of the stars, $\mathbb{E}[\phi(X_{i+1}) - \phi(X_i) \mid X_i = S] = \Theta(n^{-3})$.

However, it is possible to specialize equation (3.1) to give better bounds for restricted classes of graphs. For example, if we consider graphs of bounded degree then $(\deg x \deg y)^{-1} = \mathcal{O}(1)$ and the expected change in ϕ is $\mathcal{O}(\frac{1}{n})$.

To bound the expected absorption time, we use martingale techniques. It is well known how to bound the expected absorption time using a potential function that decreases in expectation until absorption. This has been made explicit by Hajek [11] and we use following formulation based on that of He and Yao [12]. The proof is essentially theirs but modified to give a slightly stronger result.

THEOREM 3.1. *Let $(Y_i)_{i \geqslant 0}$ be a Markov chain with state space Ω, where Y_0 is chosen from some set $I \subseteq \Omega$. If there are constants $k_1, k_2 > 0$ and a non-negative function $\psi \colon \Omega \to \mathbb{R}$ such that*

- *$\psi(S) = 0$ for some $S \in \Omega$,*

- *$\psi(S) \leqslant k_1$ for all $S \in I$ and*

- *$\mathbb{E}[\psi(Y_i) - \psi(Y_{i+1}) \mid Y_i = S] \geqslant k_2$ for all $i \geqslant 0$ and all S with $\psi(S) > 0$,*

then $\mathbb{E}[\tau] \leqslant k_1/k_2$, where $\tau = \min\{i : \psi(Y_i) = 0\}$.

Proof. By the third condition, the chain is a supermartingale so it converges to zero almost surely [25, Theorem II-2-9].

$\mathbb{E}[\psi(Y_i) \mid \psi(Y_0) > 0]$
$\quad = \mathbb{E}\big[\mathbb{E}\big[\psi(Y_{i-1}) + \big(\psi(Y_i) - \psi(Y_{i-1})\big) \mid Y_{i-1}\big]$
$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \mid \psi(Y_0) > 0\big]$
$\quad \leqslant \mathbb{E}[\psi(Y_{i-1}) - k_2 \mid \psi(Y_0) > 0].$

Induction on i gives $\mathbb{E}[\psi(Y_i) \mid \psi(Y_0) > 0] \leqslant \mathbb{E}[\psi(Y_0) - ik_2 \mid \psi(Y_0) > 0]$ and, from the definition of the stopping time τ,

$$\begin{aligned}
0 &= \mathbb{E}[\psi(Y_\tau) \mid \psi(Y_0) > 0] \\
&\leqslant \mathbb{E}[\psi(Y_0)] - k_2\mathbb{E}[\tau \mid \psi(Y_0) > 0] \\
&\leqslant k_1 - k_2\mathbb{E}[\tau \mid \psi(Y_0) > 0].
\end{aligned}$$

The possibility that $\psi(Y_0) = 0$ can only decrease the expected value of τ since, in that case, $\tau = 0$. Therefore, $\mathbb{E}[\tau] \leqslant \mathbb{E}[\tau \mid \psi(Y_0) > 0] \leqslant k_1/k_2$. $\square$

THEOREM 3.2. *Let $G = (V, E)$ be a graph of order n. For $r < 1$, the absorption time τ of the Moran process on G satisfies*

$$\mathbb{E}[\tau] \leqslant \frac{1}{1-r}n^3.$$

Proof. Let $(Y_i)_{i \geqslant 0}$ be the process on G that behaves identically to the Moran process except that, if the mutants reach fixation, we introduce a new non-mutant on a vertex chosen uniformly at random. That is, from the state V, we move to $V - x$, where x is chosen u.a.r., instead of staying in V. Writing $\tau' = \min\{i : Y_i = \emptyset\}$ for the absorption time of this new process, it is clear that $\mathbb{E}[\tau] \leqslant \mathbb{E}[\tau']$.

The function ϕ meets the criteria for ψ in the statement of Theorem 3.1 with $k_1 = 1$ and $k_2 = (1 - r)n^{-3}$. The first two conditions of the theorem are obviously satisfied. For $S \subset V$, the third condition is satisfied by Lemma 3.1 and we have

$$\mathbb{E}[\phi(Y_i) - \phi(Y_{i+1}) \mid Y_i = V] = \frac{1}{n}\sum_{x \in V}\frac{1}{\deg x} > \frac{1}{n} > k_2\,.$$

Therefore, $\mathbb{E}[\tau] \leqslant \mathbb{E}[\tau'] \leqslant \frac{1}{1-r}n^3$. $\square$

The following corollary is immediate from Markov's inequality.

COROLLARY 3.1. *The Moran process on G with fitness $r < 1$ reaches absorption within t steps with probability at least $1 - \varepsilon$, for any $\varepsilon \in (0, 1)$ and any $t \geqslant \frac{1}{1-r}n^3/\varepsilon$.*

For $r > 1$, the proof needs slight adjustment because, in this case, ϕ increases in expectation.

THEOREM 3.3. *Let $G = (V, E)$ be a graph of order n. For $r > 1$, the absorption time τ of the Moran process on G satisfies*

$$\mathbb{E}[\tau] \leqslant \frac{r}{r-1}n^3\phi(G) \leqslant \frac{r}{r-1}n^4.$$

Proof. Let $(Y_i)_{i \geqslant 0}$ be the process that behaves identically to the Moran process $(X_i)_{i \geqslant 0}$ except that, if the set of mutants is empty, a new mutant is created on a vertex chosen uniformly at random. Setting $\tau' = \min\{i : Y_i = V\}$, we have $\mathbb{E}[\tau] \leqslant \mathbb{E}[\tau']$.

Putting $\psi(S) = \phi(G) - \phi(S)$, $k_1 = \phi(G) \leqslant n$ and $k_2 = (1 - \frac{1}{r})n^{-3}$ satisfies the conditions of Theorem 3.1 — the third condition follows from Lemma 3.1 for $\emptyset \subset S \subset V$ and

$$\mathbb{E}[\psi(Y_i) - \psi(Y_{i+1}) \mid Y_i = \emptyset] = \frac{1}{n}\sum_{x \in V}\frac{1}{\deg x} > \frac{1}{n} > k_2.$$

The result follows from Theorem 3.1. $\square$

COROLLARY 3.2. *The Moran process on G with fitness $r > 1$ reaches absorption within t steps with probability at least $1 - \varepsilon$, for any $\varepsilon \in (0,1)$ and any $t \geqslant \frac{r}{r-1} n^3 \phi(G)/\varepsilon$.*

The $\mathcal{O}(n^4)$ bound in Theorem 3.3 does not seem to be very tight and could, perhaps, be improved by a more careful analysis, which we leave for future work. In simulations, we have not found any class of graphs where the expected fixation time is $\Omega(n^3)$ for $r > 1$. The graphs G_n described after Lemma 3.1 are the slowest we have found but, even on those graphs, the absorption time is, empirically, still $\mathcal{O}(n^3)$. Note that $n - 2 < \phi(G_n) < n - 1$ so, for these graphs, even the bound of $\frac{r}{r-1} n^3 \phi(G_n)$ is $\mathcal{O}(n^4)$.

The case $r = 1$ is more complicated as Lemma 3.1 no longer bounds the expected increase in ϕ away from zero. However, it still shows that the expectation is non-decreasing, which allows us to use standard martingale techniques. The proof of the following is partly adapted from the proof of Lemma 3.4 in [21].

To avoid cumbersome notation, we write $\phi_0' = \frac{1}{n} \sum_{x \in V} (\deg x)^{-2}$. For the Moran process $(X_i)_{i \geqslant 0}$ on G, this is $\mathbb{E}[\phi(X_0)^2]$.

THEOREM 3.4. *The expected absorption time for the Moran process $(X_i)_{i \geqslant 0}$ with $r = 1$ on a graph $G = (V, E)$ is at most $n^3(\phi(G)^2 - \phi_0')$.*

Proof. Let $m = \phi(G)/2$ and let $\psi_i = m - \phi(X_i)$. Thus, $-m \leqslant \psi_i \leqslant m$ for all i.

By Lemma 3.1, $\mathbb{E}[\phi(X_{i+1}) \mid X_i] \geqslant \phi(X_i)$ so

$$(3.2) \qquad \mathbb{E}[\psi_{i+1} \mid X_i] \leqslant \psi_i \,.$$

From the definition of the process, $\psi_{i+1} \neq \psi_i$ if, and only if, $X_{i+1} \neq X_i$. Therefore, $\mathbb{P}[\psi_{i+1} \neq \psi_i] = \mathbb{P}[X_{i+1} \neq X_i]$ and, for $0 < |X_i| < n$, this probability is at least n^{-2} because there is at least one edge from a mutant to a non-mutant. From the definition of ϕ, if $\psi_{i+1} \neq \psi_i$ then $|\psi_{i+1} - \psi_i| \geqslant n^{-1}$. When $|\psi_i| < m$, it follows that

$$(3.3) \qquad \mathbb{E}[(\psi_{i+1} - \psi_i)^2 \mid X_i] \geqslant n^{-3} \,.$$

Let $t_0 = \min\{t : |\psi_t| = m\}$, which is a stopping time for the sequence $(\psi_t)_{t \geqslant 0}$ and is also the least t for which $X_t = \emptyset$ or $X_t = V$. Let

$$Z_t = \begin{cases} \psi_t^2 - 2m\psi_t - n^{-3}t & \text{if } |\psi_t| < m \\ 3m^2 - n^{-3}t_0 & \text{otherwise.} \end{cases}$$

We now show that $(Z_t)_{t \geqslant 0}$ is a submartingale. This is trivial for $t \geqslant t_0$, since then we have $Z_{t+1} = Z_t$. In

the case where $t < t_0$,

$$\mathbb{E}[Z_{t+1} - Z_t \mid X_t]$$
$$\geqslant \mathbb{E}[\, \psi_{t+1}^2 - 2m\psi_{t+1} - n^{-3}(t+1)$$
$$\qquad\qquad - \psi_t^2 + 2m\psi_t + n^{-3}t \mid X_t]$$
$$= \mathbb{E}[-2m(\psi_{t+1} - \psi_t) + \psi_{t+1}^2 - \psi_t^2 - n^{-3} \mid X_t]$$
$$= \mathbb{E}[\, 2(\psi_t - m)(\psi_{t+1} - \psi_t)$$
$$\qquad\qquad + (\psi_{t+1} - \psi_t)^2 - n^{-3} \mid X_t]$$
$$\geqslant 0 \,.$$

The first inequality is because $3m^2 \geqslant \psi_t^2 - 2m\psi_t$ for all t, since $|\psi_t| \leqslant m$. The final inequality comes from equations (3.2) and (3.3). Note also that $\mathbb{E}[Z_{t+1} - Z_t \mid X_t] \leqslant 6m^2 < \infty$ in all cases.

We have

$$\mathbb{E}[Z_0] = \mathbb{E}\big[\big(m - \phi(X_0)\big)^2 - 2m\big(m - \phi(X_0)\big)\big]$$
$$= \mathbb{E}[\phi(X_0)^2 - m^2]$$
$$= \phi_0' - m^2$$

and $\mathbb{E}[Z_{t_0}] = 3m^2 - n^{-3}\mathbb{E}[t_0]$. Because t_0 is a stopping time, the optional stopping theorem says that $\mathbb{E}[Z_{t_0}] \geqslant \mathbb{E}[Z_0]$, as long as $\mathbb{E}[t_0] < \infty$, which we will show in a moment. It follows, then, that

$$3m^2 - n^{-3}\mathbb{E}[t_0] \geqslant \phi_0' - m^2,$$

which gives

$$\mathbb{E}[t_0] \leqslant n^3(4m^2 - \phi_0') = n^3(\phi(G)^2 - \phi_0') \,,$$

as required.

It remains to establish that t_0 has finite expectation. Consider a block of n successive stages $X_k, \ldots, X_{k+n-1}$. If the Moran process has not already reached absorption by X_k, then $|X_k| \geqslant 1$. Consider any sequence of reproductions by which a single mutant in X_k could spread through the whole graph. Each transition in that sequence has probability at least n^{-2} so the sequence has probability at least $p = (n^{-2})^n$, which means that the probability of absorption within the block is at least this value. But then the expected number of blocks before absorption is at most

$$\sum_{i>0} (1-p)^{i-1} = \frac{1}{1 - (1-p)} = \frac{1}{p} \,.$$

and, therefore, $\mathbb{E}[t_0] < \infty$ as required. $\square$

COROLLARY 3.3. *(i) The expected absorption time for the Moran process with $r = 1$ on any graph is at most $t = \phi(G)^2 n^3$. (ii) For any $\varepsilon \in (0, 1)$, the process reaches absorption within t/ε steps with probability at least $1 - \varepsilon$.*

Proof. The first part is immediate from the previous theorem and the fact that $\phi_0' > 0$. The second part follows by Markov's inequality. $\square$

4 Approximation algorithms

We now have all the components needed to present our fully polynomial randomized approximation schemes (FPRAS) for the problem of computing the fixation probability of a graph, where $r \geqslant 1$, and for computing the extinction probability for all $r > 0$. Recall that an FPRAS for a function f is a randomized algorithm g that, given input X, gives an output satisfying

$$(1 - \varepsilon)f(X) \leqslant g(X) \leqslant (1 + \varepsilon)f(X)$$

with probability at least $\frac{3}{4}$ and has running time polynomial in both n and $\frac{1}{\varepsilon}$. In both of the following theorems, we require that r be encoded in unary, to ensure that N is bounded by a polynomial in the size of the input.

THEOREM 4.1. *There is an FPRAS for* MORAN FIXATION, *for* $r \geqslant 1$.

Proof. The algorithm is to simulate the Moran process on G for some number T of steps (to be defined shortly), $N = \lceil \frac{1}{2}\varepsilon^{-2}n^2 \ln 16 \rceil$ times and compute the proportion of simulations that reached fixation. If any simulation has not reached absorption (fixation or extinction) after T steps, we abort and immediately return an error value. For $r > 1$, let $T = \lceil \frac{8r}{r-1}Nn^4 \rceil$ and, for $r = 1$, let $T = 8Nn^5$.

Note that each transition of the Moran process can be simulated in $\mathcal{O}(1)$ time. Maintaining lists of the mutant and non-mutant vertices allows the reproducing vertex to be chosen in constant time and storing a list of each vertex's neighbours allows the same for the vertex where the offspring is sent. Therefore, the total running time is $\mathcal{O}(NT)$ steps, which is polynomial in n and $\frac{1}{\varepsilon}$, as required.

It remains to show that the algorithm operates within the required error bounds. For $i \in \{1, \ldots, N\}$, let $X_i = 1$ if the ith simulation of the Moran process reaches fixation and $X_i = 0$ otherwise. Assuming all simulation runs reach absorption, the output of the algorithm is $p = \frac{1}{N}\sum_i X_i$. By Hoeffding's inequality and writing $f = f_{G,r}$, we have

$$\mathbb{P}[|p - f| > \varepsilon f] \leqslant 2\exp(-2\varepsilon^2 f^2 N)$$
$$= 2\exp(-2f^2 n^2 \ln 16)$$
$$\leqslant \tfrac{1}{8},$$

where the second inequality is because, by Lemma 2.1, $f \geqslant \frac{1}{n}$.

Now, the probability that any individual simulation has not reached absorption after T steps is at most $\frac{1}{8N}$ by Corollary 3.2 for $r > 1$ and Corollary 3.3 for $r = 1$. Taking a union bound, the probability of aborting and returning an error because at least one of the N simulations was cut off before reaching absorption is at most $\frac{1}{8}$. Therefore, with probability at least $\frac{3}{4}$, the algorithm returns a value within a factor of ε of $f_{G,r}$. $\square$

Note that this technique fails for disadvantageous mutants ($r < 1$) because there is no analogue of Lemma 2.1 giving a polynomial lower bound on $f_{G,r}$. As such, an exponential number of simulations may be required to achieve the desired error probability. However, we can give an FPRAS for the extinction probability for all $r > 0$. Although the extinction probability is just $1 - f_{G,r}$, there is no contradiction because a small relative error in $1 - f_{G,r}$ does not translate into a small relative error in $f_{G,r}$ when $f_{G,r}$ is, itself, small.

THEOREM 4.2. *There is an FPRAS for* MORAN EXTINCTION *for all* $r > 0$.

Proof. The algorithm is as above but taking $N = \lceil \frac{1}{2}\varepsilon^{-2}(r + n)^2 \ln 16 \rceil$ and $T = \lceil \frac{8}{1-r}Nn^3 \rceil$ for $r < 1$; we keep T as the same multiples of N as before for $r \geqslant 1$. The proof proceeds as before but using Lemma 2.2 to show that the extinction probability is at least $(r+n)^{-1}$ and Corollary 3.1 to bound the probability that a run is truncated when $r < 1$. $\square$

It remains open whether other techniques could lead to an FPRAS for MORAN FIXATION when $r < 1$.

References

[1] D. J. Aldous and J. A. Fill. Reversible Markov Chains and Random Walks on Graphs. Monograph in preparation. Available at `http://www.stat.berkeley.edu/aldous/RWG/book.html`.

[2] T. Antal and I. Scheuring. Fixation of strategies for an evolutionary game in finite populations. *Bulletin of Mathematical Biology*, 68:1923–1944, 2006.

[3] E. Berger. Dynamic monopolies of constant size. *Journal of Combinatorial Theory, Series B*, 83:191–200, 2001.

[4] M. Broom, C. Hadjichrysanthou and J. Rychtář. Evolutionary games on graphs and the speed of the evolutionary process. *Proceedings of the Royal Society A*, 466(2117):1327–1346, 2010.

[5] M. Broom, C. Hadjichrysanthou and J. Rychtář. Two results on evolutionary processes on general non-directed graphs. *Proceedings of the Royal Society A*, 466(2121):2795–2798, 2010.

[6] M. Broom and J. Rychtář. An analysis of the fixation probability of a mutant on special classes of non-directed graphs. *Proceedings of the Royal Society A*, 464(2098):2609–2627, 2008.

[7] M. Broom, J. Rychtář and B. Stadler. Evolutionary dynamics on small order graphs. *Journal of Interdisciplinary Mathematics*, 12:129–140, 2009.

[8] R. Durrett. *Lecture Notes on Particle Systems and Percolation*. Wadsworth Publishing Company, 1988.

[9] D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, 2010.

[10] H. Gintis. *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*. Princeton University Press, 2000.

[11] B. Hajek. Hitting-time and occupation-time bounds implied by drift analysis with applications. *Advances in Applied Probability*, 14(3):502–525, 1982.

[12] J. He and X. Yao. Drift analysis and average time complexity of evolutionary algorithms. *Artificial Intelligence*, 127:57–85, 2001.

[13] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.

[14] L. A. Imhof. The long-run behavior of the stochastic replicator dynamics. *Annals of Applied Probability*, 15(1B):1019–1045, 2005.

[15] M. Kandori, G. J. Mailath and R. Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):29–56, 1993.

[16] S. Karlin and H. M. Taylor. *A First Course in Stochastic Processes*. Academic Press, 2nd edition, 1975.

[17] R. M. Karp and M. Luby. Monte-Carlo algorithms for enumeration and reliability problems. In *Proceedings of 24th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 56–64, 1983.

[18] D. Kempel, J. Kleinberg and E. Tardos. Influential nodes in a diffusion model for social networks. In *Proceedings of the 32nd International Colloquium on Automata, Languages and Programming (ICALP)*, volume 3580 of *Lecture Notes in Computer Science*, pages 1127–1138. Springer, 2005.

[19] E. Lieberman, C. Hauert and M. A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433:312–316, 2005.

[20] T. M. Liggett. *Interacting Particle Systems*. Springer, 1985.

[21] M. Luby, D. Randall and A. Sinclair. Markov chain algorithms for planar lattice structures. *SIAM Journal on Computing*, 31(1):167–192, 2001.

[22] G. B. Mertzios, S. Nikoletseas, C. Raptopoulos and P. G. Spirakis. Natural models for evolution on networks. In *Proceedings of the 7th Workshop on Internet and Network Economics (WINE)*, 2011. (To appear.)

[23] P. A. P. Moran. Random processes in genetics. *Proceedings of the Cambridge Philosophical Society*, 54(1):60–71, 1958.

[24] E. Mossel and S. Roch. On the submodularity of influence in social networks. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing (STOC)*, pages 128–134, 2007.

[25] J. Neveu. *Discrete-Parameter Martingales*. North-Holland, 1975.

[26] M. A. Nowak. *Evolutionary Dynamics: Exploring the Equations of Life*. Harvard University Press, 2006.

[27] H. Ohtsuki and M. A. Nowak. Evolutionary games on cycles. *Proceedings of the Royal Society B*, 273(1598): 2249–2256, 2006.

[28] J. Rychtář and B. Stadler. Evolutionary dynamics on small-world networks. *International Journal of Computational and Mathematical Sciences*, 2(1):1–4, 2008.

[29] W. H. Sandholm. *Population Games and Evolutionary Dynamics*. MIT Press, 2011.

[30] C. Taylor, D. Fudenberg, A. Sasaki and M. A. Nowak. Evolutionary game dynamics in finite populations. *Bulletin of Mathematical Biology*, 66(6):1621–1644, 2004.

[31] C. Taylor, Y. Iwasa and M. A. Nowak. A symmetry of fixation times in evolutionary dynamics. *Journal of Theoretical Biology*, 243(2):245–251, 2006.

[32] A. Traulsen and C. Hauert. Stochastic evolutionary game dynamics. In *Reviews of Nonlinear Dynamics and Complexity*, volume 2. Wiley, 2009.